

We claim:

1. A method of recognizing a coupon comprising the steps of:

scanning the coupon to generate an electronic representation;

comparing segments of the electronic representation with a defined category of patterns, wherein any segments that match one of the patterns is eliminated as noise;

identifying connected segments within the electronic representation;

applying a barcode search to at least one of the connected segments and any additional segments proximate thereto to determine whether the at least one of the connected segments forms a portion of a barcode sequence, and if so determining the alphanumeric characters associated with the barcode sequence;

applying an optical character recognition search to at least one of the connected segments and any additional segments proximate thereto to determine whether the at least one of the connected segments forms a portion of a text

string, and if so determining the alphanumeric characters associated with the text string;

applying a table search to at least one of the connected segments to determine whether the at least one connected segments forms any portion of a table, and if so determining the boundaries and position of the table on the coupon; and

comparing the alphanumeric characters associated with the barcode sequence, the alphanumeric characters associated with the text string, and the boundaries and position of the table with a database of coupon data to determine whether the electronic representation matches a coupon type in the database of coupon data.

2. The method of claim 1, wherein the step of scanning the coupon comprises generating a black-and-white bit map divided into a grid of columns and rows so that each element of the grid is represented as either a black or a white bit and applying skew correction to the bit map.

3. The method of claim 2, wherein the step of detecting any connected segments comprises run-length encoding the electronic representation so that each row of the grid is represented by a plurality of start and end points that

represent the start and end of a continuous run of elements and comparing the start and end points of adjacent rows to determine whether any start or end points fall between the start and end points of the adjacent rows.

4. The method of claim 1, wherein the step of comparing segments of the electronic representation with a defined category of patterns further comprises eliminating the central bit of the segments when the comparison generates a match, provided that the elimination of the central bit will not disconnect otherwise connected components.

5. The method of claim 1, wherein the steps of applying a barcode search and applying an optical character recognition search together comprise creating a table of coupon data that identifies a location and value of any barcodes and character strings that are detected.

6. The method of claim 5, wherein the step of comparing the alphanumeric characters associated with the barcode sequence, the alphanumeric characters associated with the text string, and the boundaries and position of the table with a database of coupon data further comprise comparing the location and value of any barcode sequence and any character strings that are detected with a listing of vendor data that includes a unique vendor identifier and an

PATENT

approximate location, and wherein the match is detected if the location and value of the barcode sequence or the character strings match an entry in the listing of vendor data.

7. The method of claim 6, further comprising the step of determining a customer account and an account balance after determining a coupon type associated with the matching vendor, wherein the customer account and the account balance are read from the table of coupon data.

8. A method of identifying a vendor, a customer and an account balance based upon the representation of a coupon comprising the steps of:

grouping image data into a plurality of interconnected segments;

applying barcode recognition to at least one of the interconnected segments to detect any barcode character sequences, wherein the barcode character sequences are associated with a barcode type;

applying optical character recognition to at least one of the interconnected segments to determine an optical character sequence, wherein the optical character sequence is associated with an optical character type;

applying text character recognition to at least one of the interconnected segments to determine a text character sequence, wherein the text character sequence is associated with a text type;

generating a table of the at least one barcode character sequence associated with the barcode type, the at least one optical character sequence associated with the

PATENT

optical character type, and the text character sequence associated with the text type; and

comparing at least one of:

the barcode character sequence associated with the barcode type;

the optical character sequence associated with the optical character type; and

the text character sequence associated with the text type;

to a database of vendor data and determining whether both the character sequence and the type associated therewith generate a match, wherein the match determines the vendor;

determining an expected location of a customer identifier and an expected location of an account balance based upon the determined vendor; and

determining the customer identifier and the account balance based upon the expected location and the table.

9. The method of claim 8, wherein the grouping image data into a plurality of interconnected segments further comprises run length coding.

10. The method of claim 9, further comprising the step of determining a plurality of bounding boxes, wherein each bounding box defines the limits of one of the plurality of interconnected segments.

11. The method of claim 10, further comprising the step of comparing the bounding boxes to a plurality of thresholds to identify interconnected segments comprising noise and to identify interconnected segments comprising an OCR character sequence.

12. The method of claim 11, wherein the bounding box associated with an interconnected segment identifies a height and a width, and wherein the plurality of thresholds includes a noise threshold, so that an interconnected segment is identified as noise if one of the height and width associated therewith does not exceed the noise threshold.

13. The method of claim 12, wherein the plurality of thresholds further comprises an OCR height range and an OCR width range, so that an interconnected segment is identified

PATENT

as an OCR character if the height falls within the OCR height range and the width falls within the OCR width range.

1. A method for determining whether a character is an OCR character, comprising:
determining a height of the character;
determining a width of the character;
comparing the height of the character to a height range;
comparing the width of the character to a width range;
determining whether the height of the character falls within the height range and the width of the character falls within the width range;
determining whether the character is an OCR character based on the determination of whether the height of the character falls within the height range and the width of the character falls within the width range.

14. A computer system especially suitable for determining vendor, customer and account data associated with a coupon, comprising:

a scanner configured to generate an electronic representation of a coupon;

at least one data processor operationally coupled with the scanner and configured to:

compare segments of the electronic representation with a defined category of patterns so that any segments that match one of the patterns is eliminated as noise;

identify connected segments within the electronic representation;

apply a barcode search to at least one of the connected segments and any additional segments proximate thereto to determine whether the at least one of the connected segments forms a portion of a barcode sequence, and if so to determine the alphanumeric characters associated with the barcode sequence;

apply an optical character recognition search to at least one of the connected segments and any additional

segments proximate thereto to determine whether the at least one of the connected segments forms a portion of a text string, and if so to determine the alphanumeric characters associated with the text string;

apply a table search to at least one of the connected segments to determine whether the at least one connected segments forms any portion of a table, and if so to determine the boundaries and position of the table on the coupon;

compare the alphanumeric characters associated with the barcode sequence, the alphanumeric characters associated with the text string, and the boundaries and position of the table with a database of coupon data to determine whether the electronic representation matches a coupon type in the database of coupon data.

15. The computer system of claim 14, wherein the scanner is further configured to generate a black-and-white bit map divided into a grid of columns and rows so that each element of the grid is represented as either a black or a white bit and wherein the scanner is further configured to apply skew correction to the bit map.

PATENT

16. The computer system of claim 14, further comprising a memory operationally coupled with the at least one data processor and configured to store the defined set of patterns, and wherein the defined set of patterns are selected to avoid separating connected components.

17. The computer system of claim 16, wherein the memory is further configured to store the database of coupon data.